

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

5

- (A) NAME: PIERRE FABRE MEDICAMENT
- (B) STREET: 17, AVENUE JEAN MOULIN
- (C) CITY: CASTRES
- (D) STATE: FRANCE
- (E) POSTAL CODE: 81106

10 AGENT, PHARMACEUTICAL COMPOSITION AND PREPARATION PROCESS.

(iii) NUMBER OF SEQUENCES: 75

(iv) COMPUTER READABLE FORM:

15

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS

(vi) DATA OF THE LAST APPLICATION:

- (A) NUMBER OF THE APPLICATION: FR 94 04009
- (B) FILING DATE: APRIL 06 1994

Information for SEQ ID NO: 1 G2A

20 SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Thr Val Lys Thr Lys Asn Thr Thr Thr Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys
5' - ACC GIG AAA ACC AAA AAC ACC ACC ACG ACC ACC CAG ACC CAG CGG AGC AAA CGG ACC ACC ACC AAA
150Gln Arg Gln Asn Lys Pro Pro Asn Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe
CAG CTT CAG AAC AAA CGG CGG AAC AAA CGG AAC AAC GAT TTC CAT TTC GAA GIG TTC AAC TTC
171 173 176 182 186Val Pro Cys Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn
GIG CGG TCC ATC TGC AGC AAC AAC CGG ACC TGC TGG GGG ATC TGC AAA CGT ATC CGG AAC
192Lys Lys Pro Gly Lys Lys Thr Thr Lys Pro Thr Lys Lys Pro Thr Phe Lys Thr Thr Lys
AAA AAA CGG GGC AAA AAA ACC ACC AAA CGG ACC AAA AAA CGG ACC ACC TIC AAA ACC ACC AAA
213 230Lys Asp His Lys Pro Gln Thr Thr Lys Pro Lys Glu Val Pro Thr Thr Lys Pro - C
AAA GGT CAT AAA CGG CGG ACC ACC AAA CGG AAA GAA GIG CGG ACC ACC AAA CGG - 3'

Information for SEQ ID NO: 2 G2B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Thr Ala Gln Thr Lys Gly Arg Ile Thr Thr Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys
5' - ACC GCG CAG ACC AAA GGC CGT ATC ACC ACC AGC ATC CAG ACC AAC AAA CCG AGC ACC AAA

150

Ser Arg Ser Lys Asn Pro Pro Lys Pro Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe
AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG AAA GTT GAT TAC CAC TTC GAA GIG TTC AAC TTC

171

173

176

182

186

Val Pro Cys Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser
GIG CCC TCC ACC ATC TCC GGC AAC AAC CAG CCG TCC AAA AGC ATC TCC AAA ACC ATC CCG AGC

192

Asn Lys Pro Lys Lys Pro Thr Ile Lys Pro Thr Asn Lys Pro Thr Thr Lys Thr Thr Asn
AAC AAA CCG AAA AAG AAA CCG ACC ATC AAA CCG ACC AAC AAA CCG ACC AAC AAA AAC AAC AAC

230

213
Lys Arg Asp Pro Lys Thr Pro Ala Lys Met Pro Lys Glu Ile Ile Thr Asn - C
AAA CGT GAT CCG AAA ACC CCG CGG AAA ATG CCG AAG AAG GAA ATC ATC ACC AAC - 3'

Information for SEQ ID NO: 3 G2AδCys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Thr Val Lys Thr Lys Asn Thr Thr Thr Thr Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys
5' - ACC GIG AAA ACC AAA AAC ACC ACC AGC ACC CAG CCG AGC AAA CCG ACC ACC AAA

150
Gln Arg Gln Asn Lys Pro Pro Asn Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe
CGG CGT CAG AAC AAA CCG CCG AAC AAA CCG AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC
171 173 176 182 186
Val Pro Ser Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn
GTG CGG ACC ATC TCC AGC AAC AAC CCG ACC TCC TGG GCG ATC AGC AAA CGT ATC CCG AAC
192
Lys Lys Pro Gly Lys Lys Thr Thr Thr Lys Pro Thr Lys Lys Pro Thr Phe Lys Thr Thr Lys
AAA AAA CCG GGC AAA AAA ACC AGC ACC AAA CCG ACC AAA AAA CCG ACC TIC AAA ACC ACC AAA
213
Lys Asp His Lys Pro Gln Thr Thr Lys Pro Lys Glu Val Pro Thr Thr Lys Pro - C
AAA GAT CAT AAA CCG CAG ACC ACC AAA CCG AAA GAA GIG CCG ATC ACC AAA CCG - 3'

Information for SEQ ID NO: 4 G2B5Cys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130
N - Thr Ala Gln Thr Lys Gly Arg Ile Thr Thr Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys
5' - ACC CGG CGG ACC AAA GGC CGT ATC ACC ACC AGC ACC CGG ACC AAC AAA CCG ACC ACC AAA
150
Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro Asp Asp Tyr His Phe Glu Val Phe Asn Phe
AGC CGT ACC AAA AAC CCG CCG AAA AAA CCG AAA GAT GAT TAC CTC TIC GAA GIG TIC AAC TIC
171 173 176 182 186
Val Pro Ser Ser Ile Cys Gly Asn Asn Gln Ile Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser
GTG CCC ACC ATC TCC GGC AAC AAC CGG CCG TCC AAA ACC ATC ACC AAA ACC ATC CCG AGC
192
Asn Lys Pro Lys Lys Pro Thr Ile Lys Pro Thr Asn Lys Pro Thr Thr Lys Thr Thr Asn
AAC AAA CCG AAA ATG AAA CCG ACC ATC AAA CCG ACC AAC AAA CCG ACC ACC AAA ACC ACC AAC
213
Lys Arg Asp Pro Lys Thr Pro Ala Lys Met Pro Lys Lys Glu Ile Ile Thr Asn - C
AAA CGT GAT CCG AAA ACC CCG GCG AAA ATG CCG ATG AAC GAA ATC ATC ACC AAC - 3'

Information for SEQ ID NO: 5 G1ACys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys - C
5' - AGC ATC TGC ACC AAC AAC CAG CTC TGG GCG ATC TGC AAA - 3'

Information for SEQ ID NO: 6 G1BCys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys - C
5' - AGC ATC TGC CCC AAC AAC CAG CTC TGG AAA ACC ATC TGC AAA - 3'

Information for SEQ ID NO: 7 G1A

SEQUENCE TYPE: amino acids and nucleotides

15 SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys - C
5' - ACC ATC TGC AGC AAC AAC CGG ACC TGC TGG GGG ATC AGC AAA - 3'

Information for SEQ ID NO: 8 G1B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

5 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Ser Lys - C
5' - ACC ATC TGC GGC AAC AAC CGG CTG TGC AAA AGC ATC AGC AAA - 3'

Information for SEQ ID NO: 9 G1'A

SEQUENCE TYPE: amino acids

10 SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Asp Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys - C

Information for SEQ ID NO: 10 G1'B

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Asp Gly Asn Asn Gln Leu Cys Ser Ile Cys Lys - C

Information for SEQ ID NO: 11 G1'AδC

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 14 amino acids

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Asp Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys - C

Information for SEQ ID NO: 12 G1'BδC

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Asp Gly Asn Asn Gln Leu Cys Ser Ile Ser Lys - C

Information for SEQ ID NO: 13 P40

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 335 amino acids, 1005 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

P40 = N -Ala Pro Lys Asp Asn Thr Trp Tyr Ala Gly Gly Lys Leu Gly Trp Ser
GCT CGG AAA GAT AAC ACC TGG TAT GCA GGT CGT AAA CTC CCT TGG TCC
17
Gln Tyr His Asp Thr Gly Phe Tyr Gly Asn Gly Phe Gln Asn Asn Asn Gly Pro
CAG TAT CAC GAC ACC CGT TTC TAC CGT AAC GGT TTC CGG AAC AAC AAC GGT CGG
35
Thr Arg Asn Asp Gln Leu Gly Ala Phe Gly Gly Tyr Gln Val Asn Pro
ACC CGT AAC GAT CGG CTT GGT CCT CGT CGG TTC GGT GGT TAC CGG GTC AAC CCG
53
Tyr Leu Gly Phe Glu Met Gly Tyr Asp Trp Leu Gly Arg Met Ala Tyr Lys Gly
TAC CTC CGT TTC GAA ATG GGT TAT GAC TGG CGG CGT ATG GCA TAT AAA GGC
71
Ser Val Asp Asn Gly Ala Phe Lys Ala Gln Gly Val Gln Leu Thr Ala Lys Leu
ACC GTC GAC AAC GGT CCT TTC AAA CCT CGG CGC CGT CGT ACC CCT AAA CTC
89
Gly Tyr Pro Ile Thr Asp Asp Leu Asp Ile Tyr Thr Arg Leu Gly Gly Met Val
GGT TAC CGG ATC ACT GAC GAT CGT GAC ATC TAC ACC CGT CGT CGC CGC ATG GTR
107
125
His Asp Thr Gly Val Ser Pro Val Phe Ala Gly Gly Val Glu Trp Ala Val Thr
CGC GAC ACT CGC GTC TCC CCA GTC TTT GCT CGC CGC GTC GAG TGG GCT GTC ACT
143
Arg Asp Ile Ala Thr Arg Leu Glu Tyr Gln Trp Val Asn Asn Ile Gly Asp Ala
CGT GAC ATC GCT ACC CGT CGT GAA TAC CGG TGG GTC AAC AAC ATC CGC GAC CGG

161
Gly Thr Val Gly Thr Arg Pro Asp Asn Gly Met Leu Ser Leu Gly Val Ser Tyr
CGC ACT GTG CGT ACC CGT CCT GAT AAC GGC ATG CGT AGC CGC GTC GTC TCC TAC
179
Arg Phe Gly Gln Glu Asp Ala Ala Pro Val Val Ala Pro Ala Pro Ala
CGC TTC CGT CGG GAA GAT GCT GCA CGG GTC GTC CCT CGT CCT CGT CCT CGT

197
 Pro Glu Val Ala Thr Lys His Phe Thr Leu Lys Ser Asp Val Leu Phe Asn Phe
 CCG GAA GIG CCT ACC AAG CAC TTC ACC CCG AAG TCT GAC GTT CTG TTC AAC TTC
 215
 Asn Lys Ala Thr Leu Lys Pro Glu Gly Gln Gln Ala Leu Asp Gln Leu Tyr Thr
 AAC AAA GCT ACC CTC AAA CCG GAA CGT CAG CAG GCT CTC GAT CAG CTC TAC ACT
 233
 Gln Leu Ser Asn Met Asp Pro Lys Asp Gly Ser Ala Val Val Leu Gly Tyr Thr
 CAG CTC ACC AAC ATG GAT CTC AAA GAC GGT TCC GCT GTT CTC GTC GGC TAC ACC
 251
 Asp Arg Ile Gly Ser Glu Ala Tyr Asn Gln Leu Ser Glu Lys Arg Ala Gln
 GAC CCC ATC CGT TCC GAA GCT TAC AAC CAG CTC TCT GAG AAA CGT GCT CAG
 269
 Ser Val Val Asp Tyr Leu Val Ala Lys Gly Ile Pro Ala Gly Lys Ile Ser Ala
 TCC GTT GAT TAC CTC GTC GCT AAA GGC ATC CTC GCT GGC AAA ATC TCC GCT
 287
 Arg Gly Met Gly Glu Ser Asn Pro Val Thr Gly Asn Thr Cys Asp Asn Val Lys
 CCC CCC ATG GGT GAA TCC AAC CTC GTC ACT GGC AAC ACC TGT GAC AAC GIG AAA
 305
 Ala Arg Ala Ala Leu Ile Asp Cys Leu Ala Pro Asp Arg Arg Val Glu Ile Glu
 GCT CCC GCT GCA CTC ATC GAT TGC CTC GCT CCG GAT CGT GTC GAA ATC GAA
 323
 Val Lys Gly Tyr Lys Glu Val Val Thr Gln Pro Ala Gly TER - C
 GTC AAA GGC TAC AAA GAA GTC ACT CGG CCG CCG GGT TAA - 3'
 335

Information for SEQ ID NO: 14 G2AδCF

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130
 N - Thr Val Lys Thr Lys Asn Thr Thr Thr Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys
 5' - ACC GTG AAA ACC AAA AAC ACC ACC AGC ACC ACC CAG CCG AGC AAA CCG ACC ACC AAA
 150
 Gln Arg Ile Asn Lys Pro Pro Asn Lys Pro Asn Asn Asp Ser His Ser Glu Val Ser Asn Ser
 CAG CGT CAG AAC AAA CCG CCG AAC AAA CCG AAC AAC GAT TCC CAT TCC GAA GTG TCC AAC TCC
 171 173 176 182 186
 Val Pro Ser Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn
 GTG CCG AGC ATC TCC AGC AAC AAC CCG ACC TCC TGG CCG ATC AGC AAA CGT ATC CCG AAC
 192
 Lys Lys Pro Gly Lys Lys Thr Thr Lys Pro Thr Lys Lys Pro Thr Phe Lys Thr Thr Lys
 AAA AAA CCG CCC AAA AAA ACC AGC AAA AAA CCG ACC TCC AAA ACC ACC AAA

213 230
Lys Asp His Lys Pro Gln Thr Thr Lys Pro Lys Glu Val Pro Thr Thr Lys Pro - C
AAA GAT CAT AAA CCG CAG ACC ACC AAA CCG AAA GAA GTG CGG ACC ACC AAA CCG - 3'

Information for SEQ ID NO: 15 G4A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Cys Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys - C
5' - GTG CGG TGC AGC ATC TGC ACC AAC AAC CCG ACC TGC TGG CGG ATC TGC AAA - 3'

Information for SEQ ID NO: 16 G4AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys - C
5' - GTG CGG AGC AGC ATC TGC AGC AAC AAC CCG ACC TGC TGG CGG ATC AGC AAA - 3'

Information for SEQ ID NO: 17 G4B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Cys Ser Ile Cys Gly Asn Asn Gln Lys Cys Lys Ser Ile Cys Lys - C
5' -GIG CCT TCC AGC ATC TCC GGC AAC AAC CAG CTG TGC AAA AGC ATC TGC AAA - 3'

Information for SEQ ID NO: 18 G4BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Ile Cys Gly Asn Asn Gln Lys Cys Lys Ser Ile Ser Lys - C
5' -GIG CCT AGC AGC ATC TCC GGC AAC AAC CAG CTG TGC AAA AGC ATC AGC AAA - 3'

Information for SEQ ID NO: 19 G4'A

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Asp Ser Ile Asp Ser Asn Asn Pro Thr Orn Trp Ala Ile Orn Lys - C

Information for SEQ ID NO: 20 G4'AδC

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Ile Asp Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys - C

Information for SEQ ID NO: 21 G4'B

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 17 amino acids

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Asp Ser Ile Asp Gly Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys - C

Information for SEQ ID NO: 22 G4'BδC

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Ile Asp Gly Asn Asn Gln Leu Cys Ser Ile Ser Lys - C

Information for SEQ ID NO: 23 G200A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CGG ACC CGG CGG AGC AAA CGG ACC ACC AAA CGG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GTG TTC AAC TIC GIG CGG TGC AGC ATC TGC
177 182 186
Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro Gly
AGC AAC AAC CGG ACC TCC TGG CGG ATC TCC AAA CGT ATC CGG AAC AAA AAA CGG GCC
196 200
Lys Lys Thr Thr Thr - C
AAA AAA ACC AGG ACC - 3'

Information for SEQ ID NO: 24 G198A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CGG ACC CGG CGG AGC AAA CGG ACC ACC AAA CGG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GTG TTC AAC TIC GIG CGG TGC AGC ATC TCC
177 182 186
Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro Gly
AGC AAC AAC CGG ACC TCC TGG CGG ATC TCC AAA CGT ATC CGG AAC AAA AAA CGG GCC
196 198
Lys Lys Thr - C
AAA AAA ACC - 3'

Information for SEQ ID NO: 25 G196A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CGG CGG AGC AAA CGG ACC AAC AAA CGG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TIC GIG CGG CGG TGC ACC ATC TGC
177 182 186
Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro Gly
AGC AAC AAC CGG ACC TGC TGG CGG ATC TGC AAA CGT ATC CGG AAC AAA AAA CGG GGC
196
Lys - C
AAA - 3'

Information for SEQ ID NO: 26 G194A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CGG CGG AGC AAA CGG ACC AAC AAA CGG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TIC GIG CGG CGG TGC ACC ATC TGC
177 182 186 194
Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro - C
AGC AAC AAC CGG ACC TGC TGG CGG ATC TGC AAA CGT ATC CGG AAC AAA AAA CGG - 3'

Information for SEQ ID NO: 27 G192A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CAG CGG AGC AAA CGG ACC ACC AAA CAG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GTG TIC AAC TTC GTG CGG TCC AGC ATC TGC
177 182 186 192
Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys - C
AGC AAC AAC CGG ACC TGC TGG GCG ATC TCC AAA CGT ATC CGG AAC AAA - 3'

Information for SEQ ID NO: 28 G6A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CAG CGG AGC AAA CGG ACC ACC AAA CAG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC G . GTG TIC AAC TTC GTG CGG TCC AGC ATC TGC
177 182 186 190
Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro - C
AGC AAC AAC CGG ACC TGC TGG GCG ATC TCC AAA CGT ATC CGG - 3'

Information for SEQ ID NO: 29 G7A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

158 173
N - Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile
5' - AAA CCG AAC AAC GAT TTC CAT TTC GAA GCG TTC AAC TTC GIG CGG TGC AGC ATC
176 182 186 190
Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro - C
TCC AGC AAC AAC CGG ACC TCC TGG GCG ATC TGC AAA CGT ATC CGG - 3'

Information for SEQ ID NO: 30 G200AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CGG CGG AGC AAA CGG ACC ACC AAA CAG CGT CGG AAC AAA CGG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GCG TTC AAC TTC GIG CGG AGC AGC ATC TCC
177 182 186
Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys Lys Pro Gly
ACC AAC AAC CGG ACC TCC TGG GCG ATC AGC AAA CGT ATC CGG AAC AAA CGG CGC
196 200
Lys Lys Thr Thr Thr - C
AAA AAA ACC AGC ACC - 3'

Information for SEQ ID NO: 31 G198AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CGG ACC CGG CGG AGC AAA CCG ACC ACC AAA CGG CGT CGG AAC AAA CCG CGG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CGG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TTC GIG CGG AGC AGC ATC TCC
177 182 186
Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys Lys Pro Gly
ACC AAC AAC CGG ACC TCC TGG GCG ATC AGC AAA CGT ATC CGG AAC AAA AAA CGG CGC
196 198
Lys Lys Thr - C
AAA AAA ACC - 3'

Information for SEQ ID NO: 32 G196AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CGG ACC CGG CGG AGC AAA CCG ACC ACC AAA CGG CGT CGG AAC AAA CCG CGG AAC

158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CCG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TTC GIG CCG AGC ACC ATC TCC
177 182 186
Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys Lys Pro Gly
AGC AAC AAC CCG ACC TGC TGG GCG ATC AGC AAA CGT ATC CCG AAC AAA CCG GGC
196
Lys - C
AAA - 3'

Information for SEQ ID NO: 33 G194AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CAG CCG AGC AAA CCG ACC ACC AAA CAG CGT CAG AAC AAA CCG CCG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CCG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TTC GIG CCG AGC ACC ATC TGC
177 182 186 194
Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys Lys Pro - C
AGC AAC AAC CCG ACC TGC TGG GCG ATC AGC AAA CGT ATC CCG AAC AAA CCG - 3'

Information for SEQ ID NO: 34 G192AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG ACC CAG CCG AGC AAA CCG ACC ACC AAA CAG CGT CAG AAC AAA CCG CCG AAC

158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CCG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TIC GIG CCG AGC AGC ATC TCC
177 182 186 192
Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys - C
AGC AAC AAC CCG ACC TGC TGG GIG ATC AGC AAA CGT ATC CCG AAC AAA - 3'

Information for SEQ ID NO: 35 G6AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 50 amino acids, 150 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn
5' - CAG AAC CAG CCG AGC AAA CCG ACC AAC AAA CAG CGT CAG AAC AAA CCG CCG AAC
158 173 176
Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CCG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TIC GIG CCG AGC AGC ATC TCC
177 182 186 190
Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro - C
AGC AAC AAC CCG ACC TGC TGG GIG ATC AGC AAA CGT ATC CCG - 3'

Information for SEQ ID NO: 36 G7AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

158 173
N - Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile
5' - AAA CCG AAC AAC GAT TTC CAT TTC GAA GIG TIC AAC TIC GIG CCG AGC AGC ATC

176 182 186 190
Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro - C
TGC ACC AAC AAC CCG ACC TCC TGG CGG ATC AGC AAA CGT ATC CCG - 3'

Information for SEQ ID NO: 37 G200B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro
5' - AGC ACC CAG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TTC AAC TTC GIG CTC TCC AGC ATC TGC CCC AAC AAC CAG
182 186 200
Ile Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro Thr Ile - C
CIG TGC AAA ACC ATC TCC AAA ACC ATC CCG ACC AAC AAA CCG AAA AAG AAA CCG ACC ATC - 3'

Information for SEQ ID NO: 38 G198B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro
5' - AGC ACC CAG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Ile Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TTC AAC TTC GIG CTC TCC AGC ATC TGC CCC AAC AAC CAG
182 186 196
Ile Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro - C
CIG TGC AAA ACC ATC TCC AAA ACC ATC CCG AGC AAC AAA CCG AAA AAG AAA CCG - 3'

Information for SEQ ID NO: 39 G196B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CGG ACC AAC AAA CCG AGC ACC AAA AGC CGT ACC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GTG TTC AAC TTC GIG CCC TGC AGC ATC TCC GCC AAC AAC CCG
182 186 196
Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys - C
CTG TGC AAA AGC ATC TGC AAA ACC ATC CCG AGC AAC AAA CCG AAA AAG - 3'

Information for SEQ ID NO: 40 G194B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CGG ACC AAC AAA CCG AGC ACC AAA AGC CGT ACC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GTG TTC AAC TTC GIG CCC TGC AGC ATC TCC GCC AAC AAC CCG
182 186 194
Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro - C
CTG TGC AAA AGC ATC TGC AAA ACC ATC CCG AGC AAC AAA CCG - 3'

Information for SEQ ID NO: 41 G192B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 53 amino acids, 159 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CAG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TTC AAC TTC GIG CCC TCC AGC ATC TCC GGC AAC AAC CAG
182 186 192
Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn - C
CTG TGC AAA AGC ATC TCC AAA ACC ATC CCG AGC AAC - 3'

Information for SEQ ID NO: 42 G6B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CAG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TTC AAC TTC GIG CCC TCC AGC ATC TCC GGC AAC AAC CAG
182 186 190
Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro - C
CTG TGC AAA AGC ATC TCC AAA ACC ATC CCG - 3'

Information for SEQ ID NO: 43 G7B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

158 173 176
 N - Lys Pro Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly
 5' - AAA CCG AAA GAT GAT TAC CAC TTC GAA GIG TTC AAC TTC GTG CCC TCC AGC ATC TGC GGC
182 186 190
 Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro - C
 AAC AAC CAG CTG TGC AAA ACC ATC TCC AAA ACC ATC CCG - 3'

Information for SEQ ID NO: 44 G200BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
 N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
 5' - AGC ACC CGG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
 160 173 176
 Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln
 AAA GAT GAT TAC CTC TTC GAA GIG TIC AAC TTC GTG CCC AGC AGC ATC TGC GGC AAC AAC CGG
 182 186 200
 Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro Thr Ile- C
 CCG TCC AAA AGC ATC AGC AAA ACC ATC CCG AGC AAC AAA CCG AAA ARG AAA CCG ACC ATC- 3'

Information for SEQ ID NO: 45 G198BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CGG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TIC AAC TTC GIG CCC AGC AGC ATC TGC GGC AAC AAC CAG
182 186 198
Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro - C
CIG TCC AAA AGC ATC AGC AAA ACC ATC CCG AGC AAC AAA CCG AAA ARG AAA CCG - 3'

Information for SEQ ID NO: 46 G198BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CGG ACC AAC AAA CCG AGC ACC AAA AGC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TIC AAC TTC GIG CCC AGC AGC ATC TGC GGC AAC AAC CAG
182 186 196
Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro - C
CIG TCC AAA AGC ATC AGC AAA ACC ATC CCG AGC AAC AAA CCG AAA ARG AAA CCG - 3'

Information for SEQ ID NO: 47 G194B6C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CAG ACC AAC AAA CGG AGC ACC AAA ACC CGT AGC AAA AAC CGG CGG AAA AAA CCC
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TIC AAC TIC GIG CGC AGC ACC ATC TGC GGC AAC AAC CAG
182 186 194
Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn Lys Pro - C
CGG TCC AAA ACC ATC AGC AAA ACC ATC CGG AGC AAC AAA CGG - 3'

Information for SEQ ID NO: 48 G192B6C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 53 amino acids, 159 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CAG ACC AAC AAA CGG AGC ACC AAA ACC CGT AGC AAA AAC CGG CGG AAA AAA CCC
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TTC GAA GIG TIC AAC TIC GIG CGC AGC ACC ATC TGC GGC AAC AAC CAG
182 186 192
Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn - C
CGG TCC AAA ACC ATC AGC AAA ACC ATC CGG AGC AAC - 3'

Information for SEQ ID NO: 49 G6B6C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Pro
5' - AGC ACC CGG ACC AAC AAA CCG AGC ACC AAA ACC CGT AGC AAA AAC CCG CCG AAA AAA CCG
160 173 176
Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TIC GAA GIG TTC AAC TTC GIG CCC AGC AGC ATC TCC GGC AAC AAC CAG
182 186 190
Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro - C
CTG TCC AAA ACC ATC AGC AAA ACC ATC CCG - 3'

Information for SEQ ID NO: 50 G7B6C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

158 173 176
N - Lys Pro Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly
5' - AAA CCG AAA GAT GAT TAC CAC TIC GAA GIG TTC AAC TTC GIG CCC AGC AGC ATC TCC GGC
182 186 190
Asn Asn Gln Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro - C
AAC AAC CGG CCG TCC AAA AGC ATC AGC AAA ACC ATC CCG - 3'

Information for SEQ ID NO: 51 G2V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Gln Asn Arg Lys Ile Lys Gly Gln Ser Thr Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn
5' - CAA AAC AGA AAA ATC AAA CGT CAA TCA ACA CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT
150

Pro Ser Gly Ser Ile Pro Pro Glu Asn His Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr
 CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT CAA GPC CAC AAC AAC TTC CAA ACA CTC CCC TAT
 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185

Val Pro Cys Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu
 GTC CCC TGC AGT ACA TGT GAA GGT AAT CTT GCA TGC TTA TCA CTC TGC CAT ATT GAG ACC GAA
 192

Arg Ala Pro Ser Arg Ala Pro Thr Ile Thr Leu Lys Lys Thr Pro Lys Pro Lys Thr Thr Lys
 AGA GCA CCA AGC AGA GCA CCA ACA ATC ACC CTC AAA AAG ACA CCA AAA CCA AAA ACC ACA AAA
 213 220

Lys Pro Thr Lys Thr Thr Ile His His Arg Thr Ser Pro Glu Thr Lys Leu Glu - C
 ARG CCA ACC PAG ACA ACA ATC CAT CAC AGA ACC ACC CCA GAA ACC AAA CIG CAA - 3'

Information for SEQ ID NO: 52 G2VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Gln Asn Arg Lys Ile Lys Gly Gln Ser Thr Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn
 5' - CAA AAC AGA AAA ATC AAA GGT CAA TCA ACA CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT

150
Pro Ser Gly Ser Ile Pro Pro Glu Asn His Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr
CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT
171 173 176 182 186
Val Pro Ser Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu
GTT CCC ACC AGT ACA TGT GAA GGT AAT CCT GCA TGC TTA TCA CTC ACC CAT ATT GAG ACC GAA
192
Arg Ala Pro Ser Arg Ala Pro Thr Ile Thr Leu Lys Lys Thr Pro Lys Pro Lys Thr Thr Lys
AGA GCA CCA ACC AGA GCA CCA ACA ATC ACC CTC AAA AAG ACA CCA AAA CCA AAA ACC ACA AAA
213 230
Lys Pro Thr Lys Thr Thr Ile His His Arg Thr Ser Pro Glu Thr Lys Leu Gln - C
AAG CCA ACC AAG ACA ACA ATC CAT CAC AGA ACC CCA GAA ACC AAA CTG CAA - 3'

Information for SEQ ID NO: 53 G200V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CCA CCA GCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GGT CCT TGC AGT ACA TGT GAA GGT AAT CCT
182 186 200
Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro Thr Ile - C
CCA TGC TTA TCA CTC TCC CAT ATT GAG AGG GAA AGA GCA ACC AGA GCA CCA ACA ATC - 3'

Information for SEQ ID NO: 54 G198V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC TCC AGT ACA TGT GAA CGT ATT AAT CCT
182 186 196
Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro - C
GCA TCC TTA TCA CTC TGC CAT ATT GAG AGG GAA AGA GCA CCA ACC AGA GCA CCTA - 3'

Information for SEQ ID NO: 55 G196V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC TCC AGT ACA TGT GAA CGT ATT AAT CCT
182 186 196
Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg Ala Pro Ser Arg - C
GCA TCC TTA TCA CTC TGC CAT ATT GAG AGG GAA AGA GCA CCA ACC AGA GCA CCTA - 3'

Information for SEQ ID NO: 56 G194V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein
 140
 N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
 5' - CTA CCA GCG ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
 160
 173 176
 Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu
 CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTC CTC TCC AGT ACA TGT GAA GGT AAT CCT
 182 186 194
 Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg Ala Pro - C
 GCA TGC TTA TCA CTC TGC CAT ATT GAG AGG GAA AGA GCA CCA - 3'

Information for SEQ ID NO: 57 G192V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
 140
 5' - CTA CCA GGC ACA AGA AAA CCA CCA ATT ATT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
 160
 173 176
 Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu
 CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC TGC AGT ACA TGT GAA GGT AAT CIT
 182 186 192
 Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg - C
 GCA TCC TTA TCA CTC TCC CAT ATT GGG AGG GAA X - 3'

Information for SEQ ID NO: 58 G6V

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single

TOPOLOGY: linear

5 MOLECULE TYPE: protein

140
N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT CTA TCA GGA ACC ATC CCA CCA GAA AAC CAT
160
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC TCC AGT ACA TGT GAA GGT AAT CIT
182 186 190
Ala Cys Leu Ser Leu Cys His Ile Glu Thr - C
GCA TGC TTA TCA CTC TGC CAT ATT GAG AGG - 3'

Information for SEQ ID NO: 59 G7V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

10 TOPOLOGY: linear

MOLECULE TYPE: protein

158 173 176
N - Asn His Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys
5' - AAC CAT CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC TCC AGT ACA TGT
182 186 190
Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His Ile Glu Thr - C
GAA GGT AAT CIT GCA TGC TTA TCA CTC TGC CAT ATT GAG AGG - 3'

Information for SEQ ID NO: 60 G200VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA CCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC DAT GIT CCC AGC AGT ACA TGT GAA GGT AAT CIT
182 186 200
Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro Thr Ile - C
CCA TGC TTA TCA CTC AGC CAT ATT GAG AGC GAA AGA CCA CCA AGC AGA GCA CCA ACA ATC - 3'

Information for SEQ ID NO: 61 G198VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA CCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC DAT GIT CCC AGC AGT ACA TGT GAA GGT AAT CIT
182 186 196
Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro - C
CCA TGC TTA TCA CTC AGC CAT ATT GAG AGC GAA AGA CCA CCA AGC AGA GCA CCA - 3'

Information for SEQ ID NO: 62 G196VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA CCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC AGC AGT ACA TGT GAA CGT AAT CCT
182 186 196
Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro Ser Arg - C
CCA TCC TTA TCA CTC AGC CAT ATT GAG AGG GAA AGA GCA CCA AGC AGA - 3'

Information for SEQ ID NO: 63 G198VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA CCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GTT CCC AGC AGT ACA TGT GAA CGT AAT CCT
182 186 194
Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro - C
CCA TCC TTA TCA CTC AGC CAT ATT GAG AGG GAA AGA GCA CCA AGC AGA - 3'

Information for SEQ ID NO: 64 G192VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ile Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GGT CCC AGC AGT ACA TGT GAA CGT AAT CIT
182 186 192
Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg - C
GCA TGC TTA TCA CTC AGC CAT ATT GAG AGG GAA AGA - 3'

Information for SEQ ID NO: 65 G6VδC

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single

10 TOPOLOGY: linear

MOLECULE TYPE: protein

140
N - Ile Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5' - CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT CCA TCA GGA AGC ATC CCA CCA GAA AAC CAT
160 173 176
Glx Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu
CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GGT CCC AGC AGT ACA TGT GAA CGT AAT CIT
182 186 190
Ala Cys Leu Ser Leu Ser His Ile Glu Thr - C
GCA TGC TTA TCA CTC AGC CAT ATT GAG AGG - 3'

Information for SEQ ID NO: 66 G7V8C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

158 173 176
N - Asn His Gln Asp His Asn Asn Phe Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys
5' - AAC CAT CAA GAC CAC AAC AAG TTC CAA ACA CTC CCC TAT GTT CCC AGC AGT ACA TGT
182 186 190
Glu Gly Asn Leu Ala Cys Leu Ser His Ile Glu Thr - C
GAA GGT AAT CCT GCA TCC TTA TCA CTC AGC CAT ATT GCG AGG - 3'

Information for SEQ ID NO: 67 G4V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 51 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Cys Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His - C
5' - GGT CCC TCC AGT ACA TGT GAA CGT AAT CCT GCA TCC TTA TCA CTC TGC CAT - 3'

Information for SEQ ID NO: 68 G4VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 51 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His - C
5' - GTT CCC AGC AGT ACA TGT GAA GGT AAT CTT GCA TCC TTA TCA CTC AGC CAT - 3'

Information for SEQ ID NO: 69 G4'V

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 17 amino acids

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Asp Ser Thr Asp Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His - C

Information for SEQ ID NO: 70 G4'VδC

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Thr Asp Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His - C

Information for SEQ ID NO: 71 G1V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His - C
5' - AGT ACA TGT GAA GGT AAT CIT GCA TGC TTA TCA CTC TGC CAT - 3'

Information for SEQ ID NO: 72 G1VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids,

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His - C
5' - AGT ACA TGT GAA GGT AAT CIT GCA TGC TTA TCA CTC AGC CAT - 3'

Information for SEQ ID NO: 73 G1'VδC

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Thr Asp Glu Gly Asn Leu Ala Cnn Leu Ser Leu Ser His - C

Information for SEQ ID NO: 74 BB

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 219 amino acids, 657 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

1
BB = N - Lys Tyr Gly Val Ser Asp Tyr Tyr Lys Asn Leu Ile Asn Asn Ala Lys
5' - AAA TAT GGA GTC AGT GAC TAT TAC AAG AAT CTA ATC AAC AAT CCC AAA

17
Thr Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val Val Glu Ser Ala Lys Lys
ACT GTT GAA CCC GTC AAA GAC CTT CAA GCA GTC GGT GAA TCA GCG AAG AAA

35
Ala Arg Ile Ser Glu Ala Thr Asp Gly Leu Ser Asp Phe Leu Lys Ser Gln Thr
GGG CGT ATT TCA GAA GCA ACA GAT CCC TTA TCT GAT TTC TIG AAA TCA CAA ACA

53
Pro Ala Glu Asp Thr Val Lys Ser Ile Glu Leu Ala Glu Ala Lys Val Leu Ala
CCT CCT GAA GAT ACT GAT AAA TCA ATT GAA TTA GCT GAA GCT AAA GTC TCA GCT

71

Asn Arg Glu Leu Asp Lys Tyr Gly Val Ser Asp Tyr His Lys Asn Leu Ile Asn
AAC AGA GAA CIT GAC AAA TAT GGA GTA AGT GAC TAT CAC AAG AAC CTA ATC AAC

89

Asn Ala Lys Thr Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val Val Glu Ser
AAT GCC AAA ACT GTT GAA GGT GTC AAA GAC CIT CAA GCA CAA GTT GTC GAA TCA

107

Ala Lys Lys Ala Arg Ile Ser Glu Ala Thr Asp Gly Leu Ser Asp Phe Leu Lys
GGG AAG AAA CGG CGT ATT TCA GAA GCA ACA GTT GCC TTA TCT GAT TTC TIG AAA

125

Ser Gln Thr Pro Ala Glu Asp Thr Val Lys Ser Ile Glu Leu Ala Glu Ala Lys
TCA CAA ACA CCT GCT GAA GAT ACT GTT AAA TCA ATT GAA TTA GCT GAA GCT AAA

143

Val Leu Ala Asn Arg Glu Leu Asp Lys Tyr Gly Val Ser Asp Tyr Tyr Lys Asn
GTC TTA CCT AAC AGA GAA CIT GAC AAA TAT GGA GTC AGT GAC TAT TAC AAG AAC

161

Leu Ile Asn Asn Ala Lys Thr Val Glu Gly Val Lys Ala Leu Ile Asp Glu Ile
CTA ATC AAC ATT GCT AAA ACT GTT GAA GGT GTC AAA CCA CTG ATA GAT GAA ATT

179

Leu Ala Ala Leu Pro Lys Thr Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr
TTA CCT GCA TTA CCT AAG ACT GAC ACT TAC AAA TTA ATC CIT ATT GGT AAA ACA

197

Leu Lys Gly Glu Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Arg Ser Phe
TTC AAA GGC GAA ACA ACT ACT GAA GCT GTT GAT GCT GCT ACT GCA AGA TCT TTC

215

219

Asn Phe Pro Ile Leu - C
ATT TTC CCT ATC CTC - 3'

Information for SEQ ID NO: 75 δBB = fragment of BB

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 108 amino acids, 324 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

1

 $\delta\beta\beta = N$ -Lys Tyr Gly Val Ser Asp Asp Tyr His Lys Asn Leu Ile Asn Asn Ala Lys
5' - AAA TAT GGA GTC AGT GTC TAT CAC AAG AAC CTC ATC AAC ATT GTC AAA

17

Thr Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val Val Glu Ser Ala Lys Lys
ACT GTC GAA GGT GTC AAA GAC CCT CTC CCA CAA GTC GTC GAA TCA GCG AAG AAA
35Ala Arg Ile Ser Glu Ala Thr Asp Gly Leu Ser Asp Phe Leu Lys Ser Gln Thr
GGG CGT ATT TCA GAA GCA ACA GTC GTC TTA GAT TTC TTG AAA TCA CAA ACA
53Pro Ala Glu Asp Thr Val Lys Ser Ile Glu Leu Ala Glu Ala Lys Val Leu Ala
CCT GCT GAA GTC ACT GTC AAA TCA ATT GAA TTA GTC GAA GCT AAA GTC TTA GCT
71Asn Arg Glu Leu Asp Lys Tyr Gly Val Ser Asp Tyr Tyr Lys Asn Leu Ile Asn
AAC AGA GAA CCT GAC AAA TAT GGA GTC AGT GTC TAT TAC AAG AAC CTC ATC AAC
89Asn Ala Lys Thr Val Glu Gly Val Lys Ala Leu Ile Asp Glu Ile Leu Ala Ala
ATT GTC AAA ACT GTC GAA GGT GTC AAA GCA CCT ATA GAT GAA ATT TTA GCT GCA
107 108

Leu Pro - C

TTA CCT - 3'